ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Journal of Computational and Applied Mathematics 205 (2007) 584–593

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

On the convergence of Newton's method for a class of nonsmooth operators

Ioannis K. Argyros*

Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA

Received 19 December 2004; received in revised form 17 April 2006

Abstract

We provide an analog of the Newton–Kantorovich theorem for a certain class of nonsmooth operators. This class includes smooth operators as well as nonsmooth reformulations of variational inequalities. It turns out that under weaker hypotheses we can provide under the same computational cost over earlier works [S.M. Robinson, Newton's method for a class of nonsmooth functions, Set-Valued Anal. 2 (1994) 291–305] a semilocal convergence analysis with the following advantages: finer error bounds on the distances involved and a more precise information on the location of the solution. In the local case not examined in [S.M. Robinson, Newton's method for a class of nonsmooth functions, Set-Valued Anal. 2 (1994) 291–305] we can show how to enlarge the radius of convergence and also obtain finer error estimates. Numerical examples are also provided to show that in the semilocal case our results can apply where others [S.M. Robinson, Newton's method for a class of nonsmooth functions, Set-Valued Anal. 2 (1994) 291–305] fail, whereas in the local case we can obtain a larger radius of convergence than before [S.M. Robinson, Newton's method for a class of nonsmooth functions, Set-Valued Anal. 2 (1994) 291–305].

MSC: 65H10; 65G99; 47H17; 49M15

Keywords: Newton's method; Banach space; Semilocal/local convergence; PBA approximation for a nonsmooth operator; Newton–Kantorovich method/theorem

1. Introduction

In this study we are concerned with the problem of approximating a locally unique solution x^* of equation

$$F(x) = 0, \tag{1}$$

where F is a continuous operator defined on a closed subset D of a Banach space X with values in a Banach space Y.

The well known Newton's method replaces the operator whose solution is sought by an approximate operator that we hope can be solved easier. A solution of the approximation (linearization) can be used to restart the process. Under certain conditions a Newton sequence is being generated that converges to x^*. A survey of local and semilocal convergence theorems for such an approach can be found in [1–10] and the references there.

* Tel.: +1 580 5812908; fax: +1 580 5812616.
E-mail address: ioannisa@cameron.edu.

0377-0427/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2006.05.015

In particular, in the case of the famous Newton–Kantorovich theorem [10] the approximating operators are linearizations of the form $F(x) + F'(x)(z - x)$ where given x we solve for z.

Robinson extended this theorem to hold for nonsmooth operators, where the linearization is no longer available [12]. He used as an alternative a point-based-approximation which imitates the properties of linearization of Newton's method.

Here we extend Robinson's work, and in particular we show that it is possible under weaker hypotheses, and the same computational cost to provide finer error bounds on the distances involved, and a more precise information on the location of the solution x^*.

2. Convergence analysis

We need the definitions:

Definition 1. Let F be an operator from a closed subset D of a metric space (X, m) to a normed linear space Y. Operator F has a point-based approximation PBA on D if there exist an operator $A: D \times D \to Y$ and a number $\eta \geqslant 0$ such that for each u and v in D,

(a) $\|F(v) - A(u, v)\| \leqslant \frac{1}{2}\eta m(u, v)^2$ and

(b) operator $A(u, \cdot) - A(v, \cdot)$ is Lipschitzian on D with modulus $\eta m(u, v)$.

We then say operator A is a PBA for F with modulus η.

Moreover, if for x_0 is a given point in D

(c) operator $A(u, \cdot) - A(x_0, \cdot)$ is Lipschitzian on D with modulus $\eta_0 m(u, x_0)$, then we say operator A is a PBA for F with moduli (η, η_0) at the point x_0.

Note that if F has a Fréchet-derivative that is Lipschitzian on a closed, convex set D with modulus η, and X is also a normed linear space, then by choosing

$$A(u, v) = F(u) + F'(u)(v - u) \tag{2}$$

(used in Newton's method), we get by (a)

$$\|F(v) - F(u) - F'(u)(v - u)\| \leqslant \tfrac{1}{2}\eta\|v - u\|^2,$$

whereas (b) and (c) are obtained by the Lipschitzian property of F'.

As already noted in [12] the existence of a PBA does not necessarily imply differentiability.

Newton's method is applicable to any operator having a PBA A, if solving the equation

$$A(x, z) = 0$$

for z provided that x is given is possible.

We also need the definition concerning bounds for the inverses of operators:

Definition 2. Let X, D, Y, F and x_0 be as in Definition 1. Then we define the Lipschitzian number:

$$\delta(F, D) = \inf\left\{\frac{\|F(u) - F(v)\|}{m(u, v)}, u \neq v, u, v \in D\right\}.$$

Clearly, if $\delta(F, D) \neq 0$ then F is 1–1 on D.

We also define the center-Lipschitzian number:

$$\delta(F, D, x_0) = \delta_0(F, D) = \inf\left\{\frac{\|F(u) - F(x_0)\|}{m(u, x_0)}, u \neq x_0, u \in D\right\}.$$

Set $d = \delta(F, D)$ and $d_1 = \delta_0(F, D)$.

At this point we need two lemmas. The first one is the analog of the Banach Lemma on invertible operators [12, Theorem 4(2.V)]. The second one gives conditions under which operator F is 1–1.

Lemma 1. *Let X be a Banach space, D a closed subset of X and Y a normed linear space. Let F and G be functions from D into Y, G being Lipschitzian with modulus η and center-Lipschitzian with modulus η_0. Let $x_0 \in D$ with $F(x_0) = y_0$. Assume*:

(a) $U(y_0, \alpha) = \{y \in Y \mid \|y - y_0\| \leqslant \alpha\} \subseteq F(D)$,
(b) $0 \leqslant \eta < d$,
(c) $U(x_0, d_1^{-1}\alpha) \subseteq D$ *and*
(d) $\theta_0 = (1 - \eta_0 d_1^{-1})\alpha - \|G(x_0)\| \geqslant 0$.

Then the following hold:

$$U(y_0, \theta_0) \subseteq (F + G)(U(x_0, d_1^{-1}\alpha))$$

and

$$0 < d - \eta \leqslant \delta(F + G, D).$$

Proof. Let $y \in U(y_0, \theta_0)$, $x \in U(x_0, d_1^{-1}\alpha)$ and define $Ty(x) = F^{-1}(y - G(x))$. Then we can have in turn

$$\begin{aligned}\|y - G(x) - y_0\| &\leqslant \|y - y_0\| + \|G(x) - G(x_0)\| + \|G(x_0)\| \\ &\leqslant \theta_0 + \eta_0 d_1^{-1}\alpha + \|G(x_0)\| = \alpha.\end{aligned}$$

That is the set $Ty(x) \neq \emptyset$ and in particular it contains a single point because $d > 0$. The rest follows exactly as in Lemma 3.1 in [11, p. 298].

That completes the proof of Lemma 1. □

Remark 1. In general

$$\eta_0 \leqslant \eta \tag{3}$$

and

$$d \leqslant d_1 \tag{4}$$

hold and η/η_0, d_1/d can be arbitrarily large (see Example 1). If η_0 is replaced by η and d_1 by d in Lemma 1 then our result reduces to the corresponding one in [11, Lemma 3.1]. Denote θ_0 by θ if η_0 is replaced by η and d_1 by d. Then in case strict inequality holds in any of the inequalities in (3) or (4), we get

$$\theta < \theta_0,$$

which improves the corresponding Lemma 3.1 in [11], and under the same computational cost, since in practice the computation of η, d requires the computation of η_0 or d_1, respectively.

Example 1. Let $X = Y = \mathbf{R}$, $x_0 = 0$ and define function f on $D = \mathbf{R}$ by

$$f(x) = c_0 + c_1 x + c_2 \sin e^{c_3 x}, \tag{5}$$

where c_i, $i = 0, 1, 2, 3$, are given real parameters. Using (5) and Definition 2 it can easily be seen that for c_3 large and c_2 sufficiently small, η/η_0 can be arbitrarily large.

Lemma 2. *Let X and Y be normed linear spaces, and let D be a closed subset of X. For $x_0 \in D$ and $F: D \to Y$ assume*:
function $A: D \times D \to Y$ is a PBA for F on D with moduli (η, η_0) at the point x_0;
there exists $x_0 \in D$ such that

$$U(x_0, \rho) \subseteq D.$$

Then the following holds:

$$\delta(F, U(x_0, \rho)) \geqslant d - \eta_1 \rho, \tag{6}$$

where

$$d = \delta(A(x_0, \cdot), D) \quad \textit{and} \quad \eta_1 = \eta_0 + \tfrac{1}{2}\eta.$$

Moreover, if $d - \eta_1\rho > 0$ *then* F *is* 1–1 *on* $U(x_0, \rho)$.

Proof. It follows as in the proof of Lemma 2.4 in [12, p. 295] but there are some differences. Set $x = (x_1 + x_2)/2$ for x_1 and x_2 in $U(x_0, \rho)$. Then we can write:

$$F(x_1) - F(x_2) = [F(x_1) - A(x, x_1)] + [A(x, x_1) - A(x, x_2)] + [A(x, x_2) - F(x_2)]. \tag{7}$$

We will find bounds on each one of the quantities inside the brackets above.

First by (a) of Definition 1, we obtain

$$\|F(x_i) - A(x, x_i)\| \leqslant \tfrac{1}{2}\eta\|x - x_i\|^2 = \tfrac{1}{8}\eta\|x_1 - x_2\|^2. \tag{8}$$

Using the triangle inequality we get

$$\|A(x, u) - A(x, v)\| \geqslant \|A(x_0, u) - A(x_0, v)\| - \|[A(x, u) - A(x_0, u)] - [A(x, v) - A(x_0, v)]\|. \tag{9}$$

By the definition of d and (9) we get in turn

$$\begin{aligned}\delta(A(x, \cdot), D) &\geqslant \delta(A(x_0, \cdot), D) - \sup\{\|[A(x, u) - A(x_0, u)] \\ &\qquad - [A(x, v) - A(x_0, v)]\|/\|u - v\|, u \neq v, u, v \in D\} \\ &\geqslant d - \eta_0\|x - x_0\| \geqslant d - \eta_0\rho.\end{aligned} \tag{10}$$

Hence by (7), (10) and (c) of Definition 1 we have

$$\|F(x_1) - F(x_2)\| \geqslant \left[(d - \eta_0\rho) - \tfrac{1}{4}\eta\|x_1 - x_2\|\right]\|x_1 - x_2\|,$$

and for $x_1 \neq x_2$,

$$\begin{aligned}&\|F(x_1) - F(x_2)\|/\|x_1 - x_2\| \geqslant d - \eta_0\rho - \tfrac{1}{4}\eta\|x_1 - x_2\| \\ &\quad \geqslant d - \eta_0\rho - \tfrac{1}{4}\eta\|(x_1 - x_0) + (x_0 - x_2)\| \geqslant d - \eta_0\rho - \tfrac{1}{2}\eta\rho.\end{aligned}$$

That completes of the proof of Lemma 2. □

Remark 2. If equality holds in (3) and (4), then Lemma 2 reduces to Lemma 2.4 in [12]. Otherwise our result provides a wider range for $\delta(F, U(x_0, \rho))$.

In order for us to compare our Theorem 1 with the corresponding Theorem 3.2 given by S. Robinson in [4], let r_0 and d_0 be positive numbers.

It is convenient to define scalar sequences $\{t_n\}$, $\{s_n\}$ by

$$t_{n+2} = t_{n+1} + \frac{d_0^{-1}\eta(t_{n+1} - t_n)^2}{2(1 - d_0^{-1}\eta_0 t_{n+1})}, \quad t_0 = 0, \quad t_1 = r_0 \tag{11}$$

and

$$s_{n+2} = s_{n+1} + \frac{d_0^{-1}\eta(s_{n+1} - s_n)^2}{2(1 - d_0^{-1}\eta s_{n+1})}, \quad s_0 = 0, \quad s_1 = r_0. \tag{12}$$

We set

$$h_{\mathrm{A}} = d_0^{-1}\overline{\eta} r_0, \quad \overline{\eta} = \frac{\eta_0 + \eta}{2} \tag{13}$$

and

$$h_{\mathrm{K}} = d_0^{-1}\eta r_0. \tag{14}$$

Note that

$$\tfrac{1}{2}h_{\mathrm{K}} \leqslant h_{\mathrm{A}} \leqslant h_{\mathrm{K}}. \tag{15}$$

If strict inequality holds in (3) then so does in (15). It was shown in [3] that sequence $\{t_n\}$ converges monotonically to some $t^* \in (0, 2r_0]$ provided that

$$h_{\mathrm{A}} \leqslant 1. \tag{16}$$

Moreover, by the Newton–Kantorovich theorem for nonsmooth equations [4, p. 298] sequence $\{s_n\}$ converges monotonically to some s^* with

$$0 < s^* = \frac{r_0}{h_{\mathrm{K}}}[1 - (1 - 2h_{\mathrm{K}})^{1/2}] \leqslant 2r_0 \tag{17}$$

provided that the famous Newton–Kantorovich hypothesis

$$h_{\mathrm{K}} \leqslant 1 \tag{18}$$

holds.

Note also that

$$h_{\mathrm{K}} \leqslant 1 \Rightarrow h_{\mathrm{A}} \leqslant 1 \tag{19}$$

but not vice versa unless equality holds in (3).

Under hypotheses (16) and (18) we showed in [2] (for $\eta_0 < \eta$):

$$t_n < s_n \quad (n \geqslant 2), \tag{20}$$

$$t_{n+1} - t_n < s_{n+1} - s_n \quad (n \geqslant 1), \tag{21}$$

$$t^* - t_n \leqslant s^* - s_n \tag{22}$$

and

$$t^* \leqslant s^* \leqslant 2r_0. \tag{23}$$

Note that $\{s_n\}$ was essentially used as a majorizing sequence for $\{x_n\}$ in Theorem 3.2 in [12] for Newton's method.

We can state the following semilocal convergence theorem for Newton's method involving nonsmooth operators:

Theorem 1. *Let $F: D \subseteq X \to Y$ be a continuous operator, let $x_0 \in D$ and let $r_0 > 0$. Suppose that $A: D \times D \to Y$ is a PBA for F with moduli (η, η_0) at the point $x = x_0$.*

Moreover, assume:

(a) $\delta(A(x_0, \cdot), D) \geqslant d_0 > 0$;

(b) $0 < h_A \leqslant \frac{1}{2}$;

(c) *for each $y \in U(0, d_0(t^* - r_0))$ equation $A(x_0, x) = y$ has a solution x*;

(d) *the solution $S(x_0)$ of $A(x_0, S(x_0)) = 0$ satisfies $\|S(x_0) - x_0\| \leqslant r_0$; and*

(e) $U(x_0, t^*) \subseteq D$.

Then the Newton iteration defining x_{n+1} by

$$A(x_n, x_{n+1}) = 0 \quad (n \geqslant 0)$$

is well defined, remains in $U(x_0, t^*)$ for all $n \geqslant 0$ and converges to a solution $x^* \in U(x_0, t^*)$ of equation $F(x) = 0$.

Moreover, the following estimates hold for all $n \geqslant 0$:

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n \tag{24}$$

and

$$\|x_n - x^*\| \leqslant t^* - t_n. \tag{25}$$

Furthermore, if ρ given in Lemma 2 satisfies

$$t^* < \rho < d_0 \eta_1^{-1}, \tag{26}$$

then x^* is the unique solution of equation $F(x) = 0$ in $U(x_0, \rho)$.

Proof. We use Lemma 1 with quantities F, G, α, x_0 and y_0 replaced by $A(x_0, \cdot)$, $A(x_1, \cdot) - A(x_0, \cdot)$, $d_1(t^* - r_0)$, $x_1 = S(x_0)$ and 0, respectively. Hypothesis (a) of the lemma follows from the fact that $A(x_0, x) = y$ has a unique solution x_1 (since $\delta(A(x_0, \cdot), D)) \geqslant d_0 > 0$). For hypothesis (b) we have

$$\delta(A(x_1, \cdot), D) \geqslant \delta(A(x_0, \cdot), D) - \eta_0 \|x_1 - x_0\| \geqslant d_0 - \eta_0 r_0 > 0. \tag{27}$$

Condition (c) of Lemma 1 follows immediately from the above choices. The condition $h_A \leqslant \frac{1}{2}$ is equivalent to $2r_0(\eta + \eta_0) \leqslant d_0$. We have $d_1 = \delta_0(F, D) \geqslant \delta(F, D) \geqslant d_0$, and $\|G(x_1)\| \leqslant \frac{1}{2}\eta_0 r_0^2 \leqslant \frac{1}{2}\eta r_0^2$. Then condition (d) of Lemma 1 follows from the estimate

$$\begin{aligned}
\theta &\geqslant \left(1 - \frac{\eta_0 r_0}{d_1}\right) d_1(t^* - r_0) - \frac{\eta}{2} r_0^2 = (d_1 - \eta_0 r_0)(t^* - r_0) - \frac{\eta}{2} r_0^2 \\
&\geqslant (d_0 - \eta_0 r_0)(t^* - r_0) - \frac{\eta_0}{2} r_0^2 = (d_0 - \eta_0 t_1)(t^* - t_1) - \frac{\eta}{2} t_1^2 \\
&= (d_0 - \eta_0 r_0)(t^* - t_2) + (d_0 - \eta_0 t_1)(t_2 - t_1) - \frac{\eta}{2}(t_1 - t_0)^2 \\
&= (d_0 - \eta_0 r_0)(t^* - t_2) \geqslant \eta_0 r_0 (t^* - t_2) \geqslant 0.
\end{aligned} \tag{28}$$

It follows from Lemma 1 that for each $y \in U(0, t^* - \|x_1 - x_0\|)$, the equation $A(x_1, z) = y$ has a unique solution $z = x_2$ in D since $\delta(A(x_1, \cdot), D) > 0$.

We also have $A(x_0, x_1) = A(x_1, x_2) = 0$ and $A(x_1, x_1) = F(x_1)$. By Definition 2

$$\begin{aligned}
\|x_2 - x_1\| &\leqslant \delta(A(x_1, \cdot), D)^{-1} \|A(x_1, x_2) - A(x_1, x_1)\| \\
&\leqslant d_0^{-1}(1 - d_0^{-1}\eta_0 \|x_1 - x_0\|)^{-1} \|A(x_0, x_1) - F(x_1)\| \\
&\leqslant d_0^{-1}(1 - d_0^{-1}\eta_0 \|x_1 - x_0\|)^{-1} \tfrac{1}{2}\eta \|x_1 - x_0\|^2 \leqslant t_2 - t_1.
\end{aligned} \tag{29}$$

Hence, we showed

$$\|x_{n+1} - x_n\| \leqslant t_{n+1} - t_n \tag{30}$$

and

$$U(x_{n+1}, t^* - t_{n+1}) \subseteq U(x_n, t^* - t_n) \tag{31}$$

hold for $n = 0, 1$.

Moreover, for every $z \in \overline{U}(x_1, t^* - t_1)$,

$$\|z - x_0\| \leqslant \|z - x_1\| + \|x_1 - x_0\| \leqslant t^* - t_1 + t_1 = t^* - t_0$$

implies $z \in U(x_0, t^* - t_0)$. Given they hold for $n = 0, 1, \ldots, j$, then

$$\|x_{j+1} - x_0\| \leqslant \sum_{i=1}^{j+1} \|x_i - x_{i-1}\| \leqslant \sum_{i=1}^{j+1} (t_i - t_{i-1}) = t_{j+1} - t_0 = t_{j+1}.$$

The induction can be completed by simply replacing x_0, x_1 above by x_n, x_{n+1}, respectively, and noticing that $\|x_{n+1} - x_n\| \leqslant r_0$. Indeed, for the crucial condition (d) of Lemma 1 assigning F as $A(x_n, \cdot)$, G as $A(x_{n+1}, \cdot) - A(x_n, \cdot)$, α as

$d_1(t^* - t_{n+1})$ we have, since $d_1 = \delta_0(F, D) \geqslant \delta(F, D) \geqslant d_0 - \eta_0 t_n$,

$$\begin{aligned}\theta_0 &\geqslant \left[1 - \frac{\eta(t_{n+1} - t_n)}{d_1}\right] d_1(t^* - t_{n+1}) - \frac{\eta}{2}(t_{n+1} - t_n)^2 \\ &= [d_1 - \eta(t_{n+1} - t_n)](t^* - t_{n+1}) - \frac{\eta}{2}(t_{n+1} - t_n)^2 \\ &\geqslant (d_0 - \eta t_{n+1})(t^* - t_{n+1}) - \frac{\eta}{2}(t_{n+1} - t_n)^2 \\ &= (d_0 - \eta t_{n+1})(t^* - t_{n+2}) + (d_0 - \eta t_{n+1})(t_{n+2} - t_{n+1}) - \frac{\eta}{2}(t_{n+1} - t_n)^2 \\ &= (d_0 - \eta t_{n+1})(t^* - t_{n+2}) \geqslant 0. \end{aligned} \tag{32}$$

It follows by hypothesis (b) that scalar sequence $\{t_n\}$ is Cauchy (see also Remark 3(a)). From (17) and (31) it follows $\{x_n\}$ is Cauchy too in a Banach space X and as such it converges to some $x^* \in U(x_0, t^*)$ (since $U(x_0, t^*)$ is a closed set).

Furthermore, we have

$$\begin{aligned}\|F(x_{n+1})\| &= \|F(x_{n+1}) - A(x_n, x_{n+1})\| \\ &\leqslant \tfrac{1}{2}\eta\|x_{n+1} - x_n\|^2 \leqslant \tfrac{1}{2}\eta(t_{n+1} - t_n)^2. \end{aligned} \tag{33}$$

That is $\|F(x_{n+1})\|$ converges to zero as $n \to \infty$. Therefore by the continuity of F we deduce $F(x^*) = 0$. Estimate (25) follows from (24) by using standard majorization techniques [2,10]. Moreover, for the uniqueness part using Lemma 2 and hypothesis (26) we deduce that x^* is an isolated zero of F, since F is 1–1 on $U(x_0, t^*)$.

That completes the proof of Theorem 1. □

Remark 3. If equality holds in (3), then our theorem reduces to the corresponding Theorem 3.2 in [12]. Otherwise according to (19)–(23) our theorem provides under weaker conditions, and the same computational cost finer error bounds on the distances involved and a more precise information on the location of the solution x^*. Moreover, the uniqueness of the solution x^* is shown in a larger ball (see (26) and compare it with 3.7 in [12], i.e., compare (26) with $s^* < \rho < d_0\eta^{-1}$).

Note that since $\eta_0 \in [0, \eta]$ our Theorem 1 can double (at most) the applicability of Robinson's original result. Moreover, note that as the following simple example indicates, Robinson's result cannot apply where ours can:

Example 2. Let $X = Y = \mathbf{R}$, $D = [a, 2 - a]$, $a \in [0, \frac{1}{2})$, $x_0 = 1$ and define function g on D by

$$g(x) = x^3 - a. \tag{34}$$

Then Robinson's condition (18) is violated since

$$h_{\mathrm{K}} = \tfrac{4}{3}(1 - a)(2 - a) > 1 \quad \text{for all } a \in [0, \tfrac{1}{2}). \tag{35}$$

However, our corresponding condition (16)

$$h_{\mathrm{A}} = \tfrac{1}{3}(1 - a)[(3 - a) + 2(2 - a)] \leqslant 1 \tag{36}$$

holds for all $a \in [(5 - \sqrt{13})/3, \frac{1}{2})$.

Another simple numerical example involving nonsmooth functions where strict inequality holds in (3) can be given by the following:

Example 3. Let g be a single-variable function with a Lipschitz continuous derivative. Then clearly function

$$f(x) = g(x^+) + x - x^+$$

is nonsmooth. Define also function A by

$$A(x, y) = g(x^+) + g'(x^+)(y^+ - x^+) + y - y^+.$$

In view of the estimates

$$F(y) - A(x, y) = \int_0^1 [g'(x^+ + t(y^+ - x^+)) - g'(x^+)](y^+ - x^+)\,\mathrm{d}t,$$

$$[A(u, x) - A(v, x)] - [A(u, y) - A(v, y)] = [g'(u^+) - g'(v^+)](x^+ - y^+)$$

and

$$[A(u, x) - A(x_0, x)] - [A(u, y) - A(x_0, y)] = [g'(u^+) - g'(x_0^+)](x^+ - y^+),$$

we deduce that function A is a PBA for f with Lipschitz η and center Lipschitz constants η_0 the same as function g. In particular, if g, D, x_0 are as given in Example 2 and since $\eta_0 = 3 - a$, $\eta = 2(2 - a)$ (for $x_0^+ = 1$), it follows that

$$\eta_0 < \eta \quad \text{for all } a \in (0, 1).$$

Remark 4. Condition (c) of Theorem 1 can be replaced by the stronger but more practical:

(b)$'$ For each fixed x_0 and y in $U(0, d_0(t^* - r_0)) = U$ (or in $U(0, d_0 r_0)$) operator Q given by

$$Q(x) = x - y + A(x_0, x)$$

is a contraction on U and maps U into itself.

It then follows by the contraction mapping principle [1,3] that equation $A(x_0, x) = y$ has a solution x in U.

We can also provide a local convergence result for Newton's method:

Theorem 2. *Let $F: D \subseteq X \to Y$ be as in Theorem 1. Assume:*
$x^* \in D$ is a simple solution of equation $F(x) = 0$;
operator F has a point-based approximation A on D with moduli (ℓ, ℓ_0) at the point x^*.
Moreover, assume that the following hold:

(a) $\delta(A(x^*, \cdot), D) \geqslant d^* > 0$;
(b) $d^* > (\ell_0 + \frac{1}{2}\ell)r^*$;
(c) *for each $y \in U(0, d^* r^*)$ the equation $A(x_0, x) = y$ has a solution x satisfying $\|x - x^*\| \leqslant r^*$; and*
(d) $U(x^*, r^*) \subseteq D$.

Then Newton's method is well defined, remains in $U(x^*, r^*)$ for all $n \geqslant 0$ and converges to x^* provided $x_0 \in U^0(x^*, r^*)$, so that

$$\|x_{n+1} - x^*\| \leqslant \frac{\ell \|x_n - x^*\|^2}{2(d^* - \ell_0 \|x_n - x^*\|)} \quad (n \geqslant 0). \tag{37}$$

Proof. We assign F as $A(x^*, \cdot)$, G as $A(x_n, \cdot) - A(x^*, \cdot)$, and α as $d^* r^*$. Clearly conditions (a)–(c) of Lemma 1 follow from this assignment. Concerning the remaining crucial condition (d), using the hypotheses of the theorem we can have in turn

$$\begin{aligned}\theta^* &\geqslant \left(1 - \frac{\ell_0 \|x_n - x^*\|}{d^*}\right) d^* r^* - \frac{\ell}{2}\|x_n - x^*\|^2 \\ &= (d^* - \ell_0\|x_n - x^*\|) r^* - \frac{\ell}{2}\|x_n - x^*\|^2 \\ &\geqslant (d^* - \ell_0 r^*) r^* - \frac{\ell}{2}(r^*)^2 \geqslant 0.\end{aligned} \tag{38}$$

That is all hypotheses of Lemma 1 hold. Hence x_{n+1} exists.

Finally, (37) follows as inequality (29) using Definition 2 and the estimate

$$\begin{aligned}\|x_{n+1}-x^*\| &\leqslant \delta(A(x_n,\cdot),D)^{-1}\|A(x_n,x^*)-A(x^*,x^*)\| \\ &\leqslant \frac{\ell\|x_n-x^*\|^2}{2(d^*-\ell_0\|x_n-x^*\|)}.\end{aligned} \tag{39}$$

That completes the proof of Theorem 2. □

Remark 5. (a) The local convergence of Newton's method was not studied in [12]. However, if it was it would have been as in Theorem 2 with $\ell_0=\ell$. Note, however, that as in (3) in general

$$\ell_0 \leqslant \ell \tag{40}$$

holds, and ℓ/ℓ_0 can be arbitrarily large. Denote by r_{R}^* the convergence radius have been obtained by Robinson. It then follows by (b) of Theorem 2 that

$$r_{\mathrm{R}}^*=\frac{2d^*}{3\ell}\leqslant r^*=\frac{2d^*}{2\ell_0+\ell}. \tag{41}$$

Moreover, in case strict inequality holds in (40), then so does in (41). This observation is very important in computational mathematics since our approach allows a wider choice for initial guesses x_0. Moreover, the corresponding error bounds are obtained if we set $\ell=\ell_0$ in (40). Therefore our error bounds are also finer (for $\ell_0<\ell$) and under the same computational cost.

(b) If operator A is given by linearization (2), then $A(x_n,x_{n+1})=0$ is a linear equation. Thus we have Newton's method for differentiable operators [2,3,10]. Otherwise, equation $A(x_n,x_{n+1})=0$ applies to a wide class of important real life problems for which it is not a linear equation. Such problems were discussed in detail in [12].

We complete this study with a numerical example where our convergence radius r^* compares favorably with the corresponding r_{R}^* discussed in Remark 5(a).

Example 4. Let $X=Y=\mathbf{R}$, $D=U(0,1)$, $x^*=0$, and define function f on D by

$$f(x)=e^x-1 \tag{42}$$

and A given by (2). Using (42), we get $d^*=1$, $\ell_0=e-1<e=\ell$. Therefore we deduce

$$r^*<r_{\mathrm{R}}^*.$$

In particular, we get

$$r_{\mathrm{R}}^*=.245252961$$

and

$$r^*=.254028662.$$

References

[1] S. Amat, S. Busquier, V. Candela, A class of quasi-Newton generalized Steffensen methods on Banach spaces, J. Comput. Appl. Math. 149 (2002) 397–408.

[2] I.K. Argyros, Advances in the Efficiency of Computational Methods and Applications, World Scientific Publishing Company, River Edge, NJ, 2000.

[3] I.K. Argyros, An improved error analysis for Newton-like methods under generalized conditions, J. Comput. Appl. Math. 157 (1) (2003) 169–185.

[4] I.K. Argyros, On a theorem of L.V. Kantorovich concerning Newton's method, J. Comput. Appl. Math. 155 (2003) 223–230.

[5] J.M. Gutierrez, M.A. Hernandez, M.A. Salanova, Resolution of quadratic equations in Banach space, Numer. Funct. Anal. Optim. 17 (1,2) (1996) 113–121.

[6] M.A. Hernandez, Chebyshev's approximation algorithms and applications, Comput. Math. Appl. 41 (3–4) (2001) 433–445.

[7] M.A. Hernandez, M.J. Rubio, Semilocal convergence of the Secant method under mild convergence conditions of differentiability, Comput. Math. Appl. 44 (3–4) (2002) 277–285.

[8] M.A. Hernandez, M.J. Rubio, J.A. Ezquerro, Secant-like methods for solving nonlinear integral equations of the Hammerstein type, J. Comput. Appl. Math. 115 (2000) 245–254.
[9] M.A. Hernandez, M.A. Salanova, A family of Chebyshev–Halley-type methods, Internat. J. Comput. Math. 47 (1993) 56–63.
[10] L.V. Kantorovich, G.P. Akilov, Functional Analysis in Normed Spaces, Pergamon Press, Oxford, 1982.
[11] S.M. Robinson, An implicit function theorem for a class of nonsmooth functions, Math. Oper. Res. 16 (2) (1991) 292–309.
[12] S.M. Robinson, Newton's method for a class of nonsmooth functions, Set-Valued Anal. 2 (1994) 291–305.